

Binary words avoiding a pattern and marked succession rule

S. Bilotta*

D. Merlini*

E. Pergola*

R. Pinzani*

Abstract

In this paper we study the enumeration and the construction of particular binary words avoiding the pattern $1^{j+1}0^j$. By means of the theory of Riordan arrays, we solve the enumeration problem and we give a particular succession rule, called *jumping and marked succession rule*, which describes the growth of such words according to their number of ones. Moreover, the problem of associating a word to a path in the generating tree obtained by the succession rule is solved by introducing an algorithm which constructs all binary words and then kills those containing the forbidden pattern.

1 Introduction

Binary words avoiding a given pattern $\mathbf{p} = p_0 \dots p_{h-1} \in \{0, 1\}^h$ constitute a regular language and can be enumerated in terms of the number of bits 1 and 0 by using classical results (see, e.g., [9, 10, 15]). Recently, in [2, 12], this subject has been studied in relation to the theory of Riordan arrays. The concept of *Riordan array* has been introduced in 1991 by Shapiro, Getu, Woan and Woodson [16], with the aim of defining a class of infinite lower triangular arrays with properties analogous to those of the Pascal's triangle.

Riordan arrays have been studied in relation to *succession rules* and *generating trees* associated to a certain combinatorial class, according to some enumerative parameter. In particular, we use an algebraic approach (see [12]) to study the connection between proper Riordan arrays and succession rules describing the growth, according to the number of ones, of particular binary words avoiding some fixed pattern \mathbf{p} .

In Section 2, we give some basic definitions and notation related to the notions of succession rule and generating tree. In particular, we introduce the concept of *jumping and marked succession rules* (see [7, 8]) which are succession rules acting on the combinatorial objects of a class and producing sons at different levels where appear marked or non-marked labels.

In Section 3, we give necessary and sufficient conditions for the number of words counted according to the number of their zeroes and ones to be related to proper Riordan arrays.

In Section 4, by means of the theory of Riordan arrays we solve algebraically the enumeration problem, according to the number of ones. This approach enables us to obtain a jumping and marked succession rule describing the growth of such words. In particular we show that, when the forbidden pattern has a particular shape, then each row of the related Riordan array corresponds to a level of the generating tree which generates all the binary words avoiding the pattern.

We will show that it is not possible to associate to a word a path in the generating tree obtained by the succession rule. The problem is solved in Section 5, where we introduce an algorithm for constructing all binary words having a fixed number of ones and excluding those containing the forbidden pattern \mathbf{p} .

*Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Viale G.B. Morgagni 65, 50134 Firenze, Italy.

bilotta@dsi.unifi.it, donatella.merlini@unifi.it, elisa@dsi.unifi.it, pinzani@dsi.unifi.it

2 Basic definitions and notations

A *succession rule* Ω is a system constituted by an *axiom* (a) , with $a \in \mathbb{N}$, and a set of *productions* of the form:

$$(k) \rightsquigarrow (e_1(k))(e_2(k)) \dots (e_k(k)), \quad k \in \mathbb{N}, \quad e_i : \mathbb{N} \rightarrow \mathbb{N}.$$

A production constructs, for any given label (k) , its *successors* $(e_1(k)), (e_2(k)), \dots, (e_k(k))$. In most of the cases, for a succession rule Ω , we use the more compact notation:

$$\left\{ \begin{array}{l} (a) \\ (k) \rightsquigarrow (e_1(k))(e_2(k)) \dots (e_k(k)) \end{array} \right. \quad (1)$$

The rule Ω can be represented by means of a *generating tree*, that is a rooted tree whose vertices are the labels of Ω ; where (a) is the label of the root and each node labelled (k) produces k sons labelled $(e_1(k)), (e_2(k)), \dots, (e_k(k))$, respectively. As usual, the root lies at level 0, and a node lies at level n if its parent lies at level $n - 1$. If a succession rule describes the growth of a class of combinatorial objects, then a given object can be coded by the sequence of labels met from the root of the generating tree to the object itself. We refer to [4] for further details and examples.

The concept of succession rule was introduced in [6] by Chung et al. to study reduced Baxter permutations, and was later applied to the enumeration of permutations with forbidden subsequences [5, 18].

We remark that, from the above definition, a node labelled (k) has precisely k sons. In [1], a succession rule having this property is said to be *consistent*. However, we can also consider succession rules, introduced in [7], in which the value of a label does not necessarily represent the number of its sons, and this will be frequently done in the sequel.

Regular succession rules are not sufficient to handle all enumeration problems and so we consider a slight generalization called *jumping succession rule*. Roughly speaking, the idea is to consider a set of succession rules acting on the objects of a class and producing sons at different levels.

The usual notation to indicate a jumping succession rule is the following:

$$\left\{ \begin{array}{l} (a) \\ (k) \xrightarrow{1} (e_1(k))(e_2(k)) \dots (e_k(k)) \\ (k) \xrightarrow{j} (d_1(k))(d_2(k)) \dots (d_k(k)) \end{array} \right. \quad (2)$$

The generating tree associated with (2) has the property that each node labelled (k) lying at level n produces two sets of sons, the first set at level $n + 1$ and having labels $(e_1(k)), (e_2(k)), \dots, (e_k(k))$ respectively and the second one at level $n + j$, with $j > 1$, and having labels $(d_1(k)), (d_2(k)), \dots, (d_k(k))$ respectively. For example, the jumping succession rule (3) counts the number of *2-generalized Motzkin paths* and Figure 1 shows some levels of the associated generating tree. For more details about these topics, see [8].

$$\left\{ \begin{array}{l} (1) \\ (k) \xrightarrow{1} (1)(2) \dots (k-1)(k+1) \\ (k) \xrightarrow{2} (k) \end{array} \right. \quad (3)$$

Another generalization is introduced in [13], where the authors deal with *marked succession rules*. In this case the labels appearing in a succession rule can be marked or not, therefore

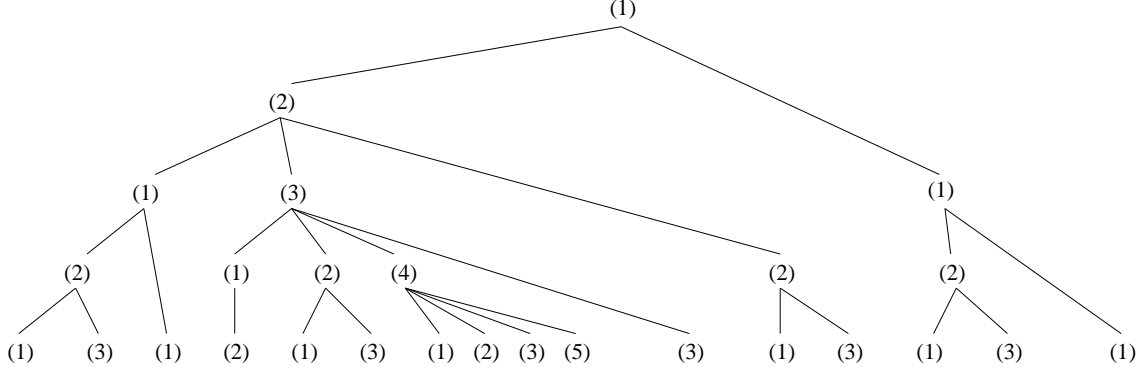


Figure 1: Four levels of the generating tree associated with the succession rule (3)

marked are considered together with usual labels. In this way a generating tree can support negative values if we consider a node labelled (\bar{k}) as opposed to a node labelled (k) lying on the same level.

A *marked generating tree* is a rooted labelled tree where appear marked or non-marked labels according to the corresponding succession rule. The main property is that, on the same level, marked labels kill or annihilate the non-marked ones with the same label value, in particular the enumeration of the combinatorial objects in a class is the difference between the number of non-marked and marked labels lying on a given level.

For any label (k) , we introduce the following notation for generating tree specifications:

$$\begin{aligned} (\bar{\bar{k}}) &= (k); \\ (k)^n &= \underbrace{(k) \dots (k)}_n, \quad n > 0. \end{aligned}$$

Each succession rule (1) can be trivially rewritten as (4)

$$\left\{ \begin{array}{l} (a) \\ (k) \rightsquigarrow (e_1(k))(e_2(k)) \dots (e_k(k))(k) \\ (k) \rightsquigarrow (\bar{k}) \end{array} \right. \quad (4)$$

For example, the classical succession rule for Catalan numbers can be rewritten in the form (5) and Figure 2 shows some levels of the associated generating tree.

$$\left\{ \begin{array}{l} (2) \\ (k) \rightsquigarrow (2)(3) \dots (k)(k+1)(k) \\ (k) \rightsquigarrow (\bar{k}) \end{array} \right. \quad (5)$$

The concept of marked labels has been implicitly used for the first time in [14], then in [7] in relation with the introduction of the signed ECO-systems. In Section 4, we show how marked succession rules appear in the enumeration of a class of particular binary words according to the number of ones. Let $F \subset \{0, 1\}^*$ be the class of binary words w such that $|w|_0 \leq |w|_1$ for any $w \in F$, $|w|_0$ and $|w|_1$ are the number of zeroes and ones in w , respectively.

In this paper we are interested in studying the subclass $F^{[p]}$ of F of binary words excluding a given pattern $\mathbf{p} = p_0 \dots p_{h-1} \in \{0, 1\}^h$, i.e. the word $w \in F^{[p]}$ that does not admit a sequence of consecutive indices $i, i+1, \dots, i+h-1$ such that $w_i w_{i+1} \dots w_{i+h-1} = p_0 p_1 \dots p_{h-1}$. Each

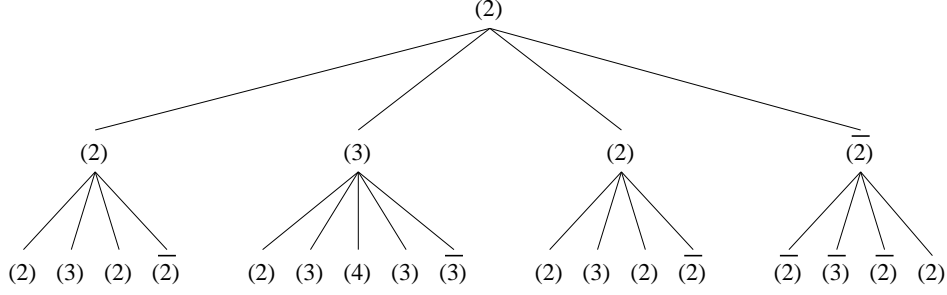


Figure 2: Three levels of the generating tree associated with the succession rule (5)

word $w \in F$ can be naturally represented as a lattice path on the Cartesian plane by associating a *rise step*, defined by $(1, 1)$ and denoted by x , to each 1's in F , and a *fall step*, defined by $(1, -1)$ and denoted by \bar{x} , to each 0's in F . From now on, we refer interchangeably to words or their graphical representations on the Cartesian plane, that is paths.

3 Binary words avoiding a pattern and Riordan arrays

In this section, we establish necessary and sufficient conditions for the number of words counted according to the number of zeroes and ones to be related to proper Riordan arrays.

This problem is interesting in the context of the Riordan arrays theory because the matrices arising there are naturally defined by recurrence relations following the characterization given in [11] (see formula (8) below). In particular, if $F_{n,k}^{[\mathbf{p}]}$ denotes the number of words excluding the pattern and having n bits 1 and k bits 0, then by using the results in [2] we have

$$F^{[\mathbf{p}]}(x, y) = \sum_{n,k \geq 0} F_{n,k}^{[\mathbf{p}]} x^n y^k = \frac{C^{[\mathbf{p}]}(x, y)}{(1 - x - y)C^{[\mathbf{p}]}(x, y) + x^{|\mathbf{p}|_1} y^{|\mathbf{p}|_0}}, \quad (6)$$

where $|\mathbf{p}|_1$ and $|\mathbf{p}|_0$ correspond to the number of ones and zeroes in the pattern and $C^{[\mathbf{p}]}(x, y)$ is the autocorrelation polynomial with coefficients given by the autocorrelation vector (see also [9, 10, 15]). For a given \mathbf{p} , this vector of bits $c = (c_0, \dots, c_{h-1})$ can be defined in terms of Iverson's bracket notation (for a predicate P , the expression $\llbracket P \rrbracket$ has value 1 if P is true and 0 otherwise) as follows: $c_i = \llbracket p_0 p_1 \dots p_{h-1-i} = p_i p_{i+1} \dots p_{h-1} \rrbracket$. In other words, the bit c_i is determined by shifting \mathbf{p} right by i positions and setting $c_i = 1$ iff the remaining letters match the original. For example, when $\mathbf{p} = 101010$ the autocorrelation vector is $c = (1, 0, 1, 0, 1, 0)$, as illustrated in Table 1, and $C^{[\mathbf{p}]}(x, y) = 1 + xy + x^2 y^2$, that is, we mark with $x^j y^i$ the tails of the pattern with j bits 1, i bits 0 and $c_{j+i} = 1$. Therefore, in this case we have:

$$F^{[\mathbf{p}]}(x, y) = \frac{1 + xy + x^2 y^2}{(1 - x - y)(1 + xy + x^2 y^2) + x^3 y^3}.$$

As another example, when $\mathbf{p} = 11100$ then $C^{[\mathbf{p}]}(x, y) = 1$ and $F^{[\mathbf{p}]}(x, y) = 1/(1 - x - y + x^3 y^2)$.

In order to study the binary words avoiding a pattern in terms of Riordan arrays, we consider the array $\mathcal{R}^{[\mathbf{p}]} = (R_{n,k}^{[\mathbf{p}]})$ given by the lower triangular part of the array $\mathcal{F}^{[\mathbf{p}]} = (F_{n,k}^{[\mathbf{p}]})$, that is, $R_{n,k}^{[\mathbf{p}]} = F_{n,n-k}^{[\mathbf{p}]}$ with $k \leq n$. More precisely, $R_{n,k}^{[\mathbf{p}]}$ counts the number of words avoiding \mathbf{p} and having length $2n - k$, n bits one and $n - k$ bits zero. Given a pattern $\mathbf{p} = p_0 \dots p_{h-1} \in \{0, 1\}^h$, let $\bar{\mathbf{p}} = \bar{p}_0 \dots \bar{p}_{h-1}$ be the pattern with $\bar{p}_i = 1 - p_i, \forall i = 0, \dots, h - 1$. We obviously have

1	0	1	0	1	0	Tails			
1	0	1	0	1	0				1
	1	0	1	0	1	0			0
		1	0	1	0	1	0		1
			1	0	1	0	1	0	0
				1	0	1	0	1	0
					1	0	1	0	0

Table 1: The autocorrelation vector for $\mathbf{p} = 101010$

n/k	0	1	2	3	4	5	6	7
0	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8
2	1	3	6	10	15	21	28	36
3	1	4	9	18	32	52	79	114
4	1	5	13	29	58	106	180	288
5	1	6	18	44	96	192	357	624
6	1	7	24	64	151	325	650	1222
7	1	8	31	90	228	524	1116	2232

Table 2: The matrix $\mathcal{F}^{[\mathbf{p}]}$ for $\mathbf{p} = 11100$

$R_{n,k}^{[\bar{\mathbf{p}}]} = F_{n,n-k}^{[\bar{\mathbf{p}}]} = F_{n-k,n}^{[\mathbf{p}]}$, therefore, the matrices $\mathcal{R}^{[\mathbf{p}]}$ and $\mathcal{R}^{[\bar{\mathbf{p}}]}$ represent the lower and upper triangular part of the array $\mathcal{F}^{[\mathbf{p}]}$, respectively. Moreover, we have $R_{n,0}^{[\mathbf{p}]} = R_{n,0}^{[\bar{\mathbf{p}}]} = F_{n,n}^{[\mathbf{p}]}$, $\forall n \in \mathbb{N}$, that is, columns zero of $\mathcal{R}^{[\mathbf{p}]}$ and $\mathcal{R}^{[\bar{\mathbf{p}}]}$ correspond to the main diagonal of $\mathcal{F}^{[\mathbf{p}]}$. Tables 2, 3 and 4 illustrate some rows for the matrices $\mathcal{F}^{[\mathbf{p}]}$, $\mathcal{R}^{[\mathbf{p}]}$ and $\mathcal{R}^{[\bar{\mathbf{p}}]}$ when $\mathbf{p} = 11100$.

We briefly recall that a Riordan array is an infinite lower triangular array $(d_{n,k})_{n,k \in \mathbb{N}}$, defined by a pair of formal power series $(d(t), h(t))$, such that $d(0) \neq 0, h(0) = 0, h'(0) \neq 0$ and the generic element $d_{n,k}$ is the n -th coefficient in the series $d(t)h(t)^k$, i.e.:

$$d_{n,k} = [t^n]d(t)h(t)^k, \quad n, k \geq 0.$$

From this definition we have $d_{n,k} = 0$ for $k > n$. An alternative definition is in terms of the so-called A -sequence and Z -sequence, with generating functions $A(t)$ and $Z(t)$ satisfying the

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	9	4	1				
4	58	29	13	5	1			
5	192	96	44	18	6	1		
6	650	325	151	64	24	7	1	
7	2232	1116	524	228	90	31	8	1

Table 3: The triangle $\mathcal{R}^{[\mathbf{p}]}$ for $\mathbf{p} = 11100$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	10	4	1				
4	58	32	15	5	1			
5	192	106	52	21	6	1		
6	650	357	180	79	28	7	1	
7	2232	1222	624	288	114	36	8	1

Table 4: The triangle $\mathcal{R}^{[\bar{\mathbf{p}}]}$ for $\bar{\mathbf{p}} = 00011$

relations:

$$h(t) = tA(h(t)), \quad d(t) = \frac{d_0}{1 - tZ(h(t))} \quad \text{with} \quad d_0 = d(0).$$

In other words, Riordan arrays correspond to matrices where each element $d_{n,k}$ is described by a linear combination of the elements in the previous row, starting from the previous column, with coefficients in A :

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \dots \quad (7)$$

Another characterization (see [11]) states that a lower triangular array $(d_{n,k})_{n,k \in \mathbb{N}}$ is Riordan if and only if there exists another array $(\alpha_{i,j})_{i,j \in \mathbb{N}}$, with $\alpha_{0,0} \neq 0$, and a sequence $(\rho_j)_{j \in \mathbb{N}}$ such that:

$$d_{n+1,k+1} = \sum_{i \geq 0} \sum_{j \geq 0} \alpha_{i,j} d_{n-i,k+j} + \sum_{j \geq 0} \rho_j d_{n+1,k+j+2}. \quad (8)$$

Matrix $(\alpha_{i,j})_{i,j \in \mathbb{N}}$ is called the A -matrix of the Riordan array. If $P^{[0]}(t), P^{[1]}(t), P^{[2]}(t), \dots$ denote the generating functions of rows $0, 1, 2, \dots$ in the A -matrix, i.e.:

$$P^{[i]}(t) = \alpha_{i,0} + \alpha_{i,1}t + \alpha_{i,2}t^2 + \alpha_{i,3}t^3 + \dots$$

and $Q(t)$ is the generating function for the sequence $(\rho_j)_{j \in \mathbb{N}}$, then we have:

$$\frac{h(t)}{t} = \sum_{i \geq 0} t^i P^{[i]}(h(t)) + \frac{h(t)^2}{t} Q(h(t)), \quad (9)$$

$$A(t) = \sum_{i \geq 0} t^i A(t)^{-i} P^{[i]}(t) + tA(t)Q(t). \quad (10)$$

The theory of Riordan arrays and the proofs of their properties can be found in [11]. The Riordan arrays which arise in the context of pattern avoidance (see [2, 12]) have the nice property to be defined by a quite simple recurrence relation following the characterization (8), while the relation induced by the A -sequence is, in general, more complex. From a combinatorial point of view, this means that it is very challenging to find a construction allowing to build objects of size $n + 1$ from objects of size n . Instead, the existence of a simple A -matrix corresponds to a possible construction from objects of different sizes less than $n + 1$. On the other hand, as we will see in Section 4, the recurrence following characterization (8) contains negative coefficients and therefore gives rise to interesting non trivial combinatorial problems.

In this paper we examine in particular the family of patterns $\mathbf{p} = 1^{j+1}0^j$ and show that the corresponding recurrence relation can be combinatorially interpreted. To this purpose, we translate the recurrence into a succession rule, as it is typically done from problems related to Riordan arrays (see, e.g., [3, 14]), and give a construction for the class of binary words avoiding the pattern \mathbf{p} .

4 The Riordan array for the pattern $\mathbf{p} = 1^{j+1}0^j$

Let us consider the family of patterns $\mathbf{p} = 1^{j+1}0^j$ and let $F_{n,k}^{[\mathbf{p}]}$ denote the number of words excluding the pattern and having n bits 1 and k bits 0; from (6) we have

$$F^{[\mathbf{p}]}(x, y) = \sum_{n,k \geq 0} F_{n,k}^{[\mathbf{p}]} x^n y^k = \frac{1}{1 - x - y + x^{j+1}y^j}. \quad (11)$$

Now, let $R_{n,k}^{[\mathbf{p}]}$ count the number of words avoiding \mathbf{p} and having n bits one and $n - k$ bits zero. Obviously we have $R_{n,k}^{[\mathbf{p}]} = F_{n,n-k}^{[\mathbf{p}]}$ with $k \leq n$. By extracting the coefficients from (11) we have:

$$[x^{n+1}y^{k+1}](1 - x - y + x^{j+1}y^j)F^{[\mathbf{p}]}(x, y) = F_{n+1,k+1}^{[\mathbf{p}]} - F_{n,k+1}^{[\mathbf{p}]} - F_{n+1,k}^{[\mathbf{p}]} + F_{n-j,k+1-j}^{[\mathbf{p}]} = 0$$

and therefore:

$$R_{n+1,k+1}^{[\mathbf{p}]} = R_{n,k}^{[\mathbf{p}]} + R_{n+1,k+2}^{[\mathbf{p}]} - R_{n-j,k}^{[\mathbf{p}]} \quad (12)$$

This is a recurrence relation of type (8) and therefore $\mathcal{R}^{[\mathbf{p}]} = (R_{n,k}^{[\mathbf{p}]})$ is a Riordan array. In particular, the coefficients of the relation correspond to $P^{[j]}(t) = -1$, $P^{[0]}(t) = 1$, and $Q(t) = 1$, therefore we have

$$\frac{h^{[\mathbf{p}]}(t)}{t} = \sum_{i \geq 0} t^i P^{[i]}(h^{[\mathbf{p}]}(t)) + \frac{h^{[\mathbf{p}]}(t)^2}{t} Q(h(t)) = 1 - t^j + \frac{h^{[\mathbf{p}]}(t)^2}{t}$$

that is,

$$h^{[\mathbf{p}]}(t)^2 - h^{[\mathbf{p}]}(t) + t - t^{j+1} = 0, \quad h^{[\mathbf{p}]}(t) = \frac{1 - \sqrt{1 - 4t + 4t^{j+1}}}{2}.$$

We explicitly observe that from formula (10) the generating function $A(t)$ of the A -sequence is the solution of a $j + 1$ degree equation $(1 - t)A(t)^{j+1} - A(t)^j + t^j = 0$. For example, when $\mathbf{p} = 11100$ by developing into series we find:

$$A(t) = 1 + t + 2t^3 - t^4 + 7t^5 - 12t^6 + 38t^7 - 99t^8 + 281t^9 + O(t^{10})$$

and this result excludes that there might exist a simple dependence of the elements in row $n + 1$ from the elements in row n . For what concerns $d^{[\mathbf{p}]}(t)$, we simply use the Cauchy formula for finding the main diagonal of matrix $\mathcal{F}^{[\mathbf{p}]}$ (see, e.g., [17, Cap. 6, p. 182]):

$$d^{[\mathbf{p}]}(t) = [x^0]F^{[\mathbf{p}]}(x, \frac{t}{x}) = \frac{1}{2\pi i} \oint F^{[\mathbf{p}]}(x, \frac{t}{x}) \frac{dx}{x}.$$

We have:

$$\frac{1}{x}F^{[\mathbf{p}]}(x, \frac{t}{x}) = \frac{-1}{x^2(1 - t^j) - x + t}$$

and in order to compute the integral, it is necessary to find the singularities $x(t)$ such that $x(t) \rightarrow 0$ with $t \rightarrow 0$ and apply the Residue theorem. In this case the right singularity is:

$$x(t) = \frac{1 - \sqrt{1 - 4t(1 - t^j)}}{2(1 - t^j)}$$

and finally we have:

$$d^{[p]}(t) = \lim_{x \rightarrow x(t)} \frac{-1}{x^2(1 - t^j) - x + t} (x - x(t)) = \frac{1}{\sqrt{1 - 4t + 4t^{j+1}}}.$$

Observe also that:

$$\frac{d^{[p]}(t) - 1}{d^{[p]}(t)h^{[p]}(t)} = 2$$

and therefore $R_{n+1,0}^{[p]} = 2R_{n+1,1}^{[p]}$. Recurrence (12) is quite simple, however, the presence of negative coefficients leads to a possible non trivial combinatorial interpretation. In order to study this problem we proceed as follows. The dependence of $R_{n+1,k+1}^{[p]}$ from the same row $n+1$ can be simply eliminated and we have:

$$\begin{aligned} R_{n+1,k+1}^{[p]} &= R_{n,k}^{[p]} - R_{n-j,k}^{[p]} + R_{n+1,k+2}^{[p]} = \\ &= R_{n,k}^{[p]} - R_{n-j,k}^{[p]} + R_{n,k+1}^{[p]} - R_{n-j,k+1}^{[p]} + R_{n+1,k+3}^{[p]} = \dots = \\ &= (R_{n,k}^{[p]} + R_{n,k+1}^{[p]} + R_{n,k+2}^{[p]} + \dots) - (R_{n-j,k}^{[p]} + R_{n-j,k+1}^{[p]} + R_{n-j,k+2}^{[p]} + \dots) \end{aligned} \quad (13)$$

Similarly we have:

$$R_{n+1,0}^{[p]} = 2(R_{n,0}^{[p]} + R_{n,1}^{[p]} + R_{n,2}^{[p]} + \dots) - 2(R_{n-j,0}^{[p]} + R_{n-j,1}^{[p]} + R_{n-j,2}^{[p]} + \dots) \quad (14)$$

Finally, by using the results in [2, 3], recurrences (13) and (14) translate into the following succession rule:

$$\left\{ \begin{array}{l} (0) \\ (k) \xrightarrow{1} (0_1)(0_2)(1) \dots (k+1) \\ (k) \xrightarrow{j+1} (\overline{0_1})(\overline{0_2})(\overline{1}) \dots (\overline{k+1}) \end{array} \right. \quad (15)$$

This rule can be represented as a tree having its root labelled (0) and where each node with label (k) at a given level n has $k+3$ sons at level $n+1$ labelled $(0_1), (0_2), (1), \dots, (k+1)$ and $k+3$ sons at level $n+j+1$ with labels $(\overline{0_1}), (\overline{0_2}), (\overline{1}), \dots, (\overline{k+1})$ (this kind of trees are called *level generating trees* in [3]). If we denote by $d_{n,k}$ the number of nodes having label k at level n in the tree and count as negative the marked nodes then we obtain matrix $\mathcal{R}^{[p]} = (R_{n,k}^{[p]})_{n,k \in \mathbb{N}}$, that is, $\mathcal{R}^{[p]}$ corresponds to the matrix associated to the rule (15). The relations between Riordan arrays and succession rules has been widely studied and we refer the reader to [3, 13, 14] for more details. We just conclude this section by observing that by using the results in [2, 12] it can be proved that the matrix $\mathcal{R}^{[\bar{p}]}$ corresponding to the pattern $\bar{p} = 0^{j+1}1^j$ is also a Riordan array.

5 A construction for the class $F^{[p]}$

In this section we define an algorithm which associates a lattice path in $F^{[p]}$, where $\mathbf{p} = x^{j+1}\bar{x}^j = 1^{j+1}0^j$, to a sequence of labels obtained by means of the succession rule (15). This give a construction for the set $F^{[p]}$ according to the number of rise steps or equivalently the number of ones.

The axiom (0) is associated to the empty path ε .

A lattice path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k , provides $k+3$ lattice paths with $n+1$ rise steps, according to the first production of (15) having $0, 0, 1, \dots, k+1$ as endpoint ordinate, respectively. The last $k+2$ labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 3). Each lattice path so obtained has the property that its rightmost suffix beginning from the x -axis, either remains strictly above the x -axis itself or ends on the x -axis by a fall step. Note that in this way the paths ending on the x -axis and having a rise step as last step are never obtained. These paths are bound to the label (0_1) of the first production in (15) and the way to obtain them will be described later in this section.

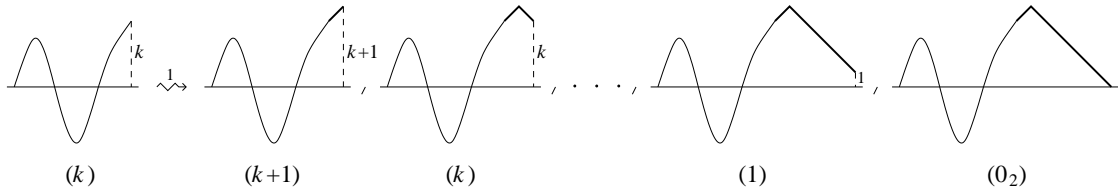


Figure 3: The mapping associated to $(k) \rightsquigarrow (0_2)(1) \dots (k+1)$ of (15)

We define a *marked forbidden pattern* \mathbf{p} as a pattern $\mathbf{p} = x^{j+1}\bar{x}^j$ whose steps cannot be divided, they must lie always in that defined sequence. Therefore, a cut operation is not possible within a marked forbidden pattern \mathbf{p} . We denote a marked forbidden pattern by marking its peak. We say that a point is strictly contained in a marked forbidden pattern if it is between two steps of the pattern itself.

A lattice path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k , provides $k+3$ lattice paths, with $n+j+1$ rise steps, according to the second production of (15) having $0, 0, 1, \dots, k+1$ as endpoint ordinate, respectively. The last $k+2$ labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathbf{p} = x^{j+1}\bar{x}^j$ followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 4). Each lattice path so obtained has the property that its rightmost suffix beginning from the x -axis, either remains strictly above the x -axis itself or ends on the x -axis by a fall step. At this point the label (0_1) due to the first and the second production of (15) yield lattice paths which either do not contain marked forbidden patterns in its rightmost suffix and end on the x -axis by a rise step or having the rightmost marked point with ordinate less than or equal to j .

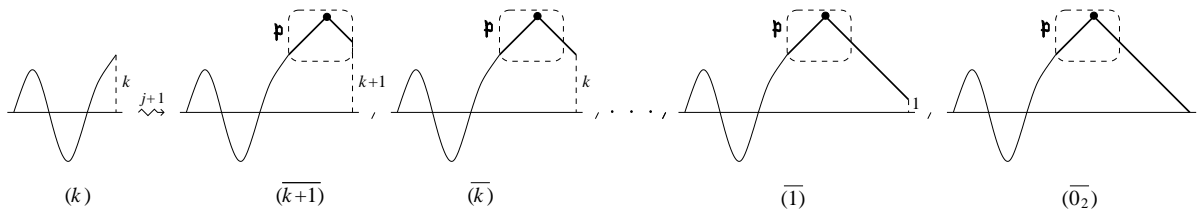


Figure 4: The mapping associated to $(k) \rightsquigarrow^{j+1} (\overline{0_2})(\overline{1}) \dots (\overline{k+1})$ of (15)

In order to obtain the label (0_1) according to the first production of (15), we consider the lattice path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps, while in order to obtain the label (0_1) according to the second production of (15), we consider the lattice path ω' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathbf{p} = x^{j+1}\bar{x}^j$ followed by k fall steps. By applying the previous actions, a path w' can be written as $w' = v\varphi$, where φ is the rightmost suffix in w' beginning from the x -axis and strictly remaining above the x -axis (see Figure 5).

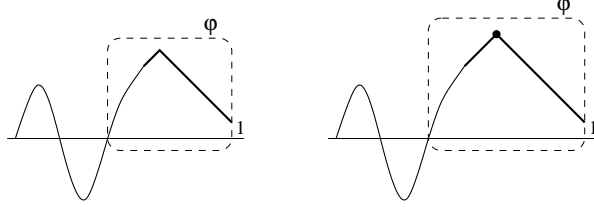


Figure 5: A graphical representation of the suffix φ in ω'

We distinguish two cases: in the first one φ does not contain any marked point and in the second one φ contains at least one marked point.

If the suffix φ does not contain any marked point, then the desired label (0_1) is associated to the path $v\varphi^c x$, where φ^c is the path obtained from φ by switching rise and fall steps (see Figure 6).

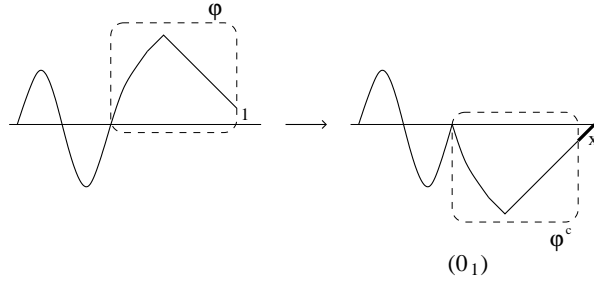


Figure 6: A graphical representation of the actions giving the label (0_1) in case of no marked points in φ

If the suffix φ contains marked points, let r be the rightmost marked point in φ having highest ordinate and t be the nearest point on the right of the marked forbidden pattern containing r with highest ordinate and which is not strictly within a marked forbidden pattern. We consider the straight line s through the point t and the leftmost point z in φ with highest ordinate, which lies above or on the line s and which is not strictly within a marked forbidden pattern (see the left side of Figure 7.a)). Obviously, if the straight line s does not intersect any points on the left of t (see the left side of Figure 7.b)) or intersects only points lying strictly within a marked forbidden pattern (see the left side of Figure 7.c)), then $z \equiv t$.

The desired label (0_1) is associated to the path obtained by concatenating a fall step \bar{x} with the path in φ running from z to the endpoint of the path, say α , and the path running from the initial point in φ to z , say β (see Figure 7 and 8).

This last mapping can be inverted as follows. Let d be the rightmost fall step in a path ω' labelled (0_1) such that it begins from the x -axis and each marked point, on its right, has ordinate less than or equal to j . Let $\omega' = \omega d\varphi'$ and l the rightmost point in φ' with lowest

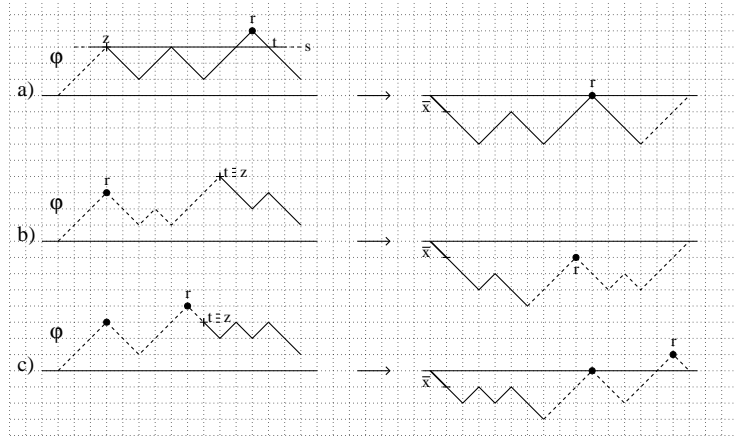


Figure 7: Some examples of the actions giving the label (0_1) in the case of marked points in φ , $\mathbf{p} = x^2\bar{x}$

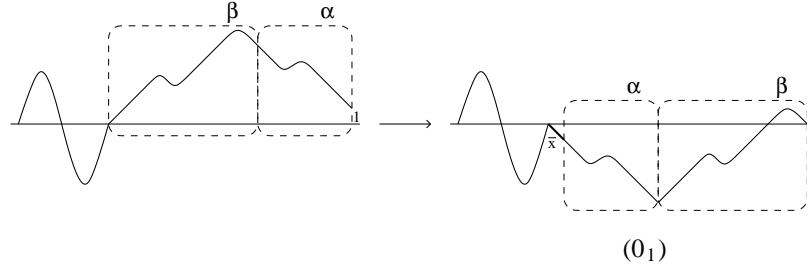


Figure 8: A graphical representation of the cut and paste actions giving the label (0_1) in case of marked points in φ

ordinate. The inverted lattice path of ω' is given by $\omega\beta\alpha$, where β is the path in φ' running from l to the endpoint of the path and α is the path running from the initial point in φ' to l (see Figure 9).

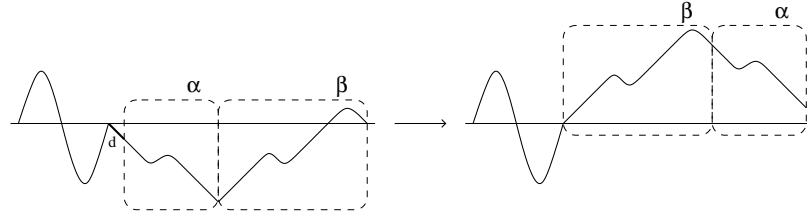


Figure 9: A graphical representation of the lattice path obtained by means of the inverted mapping related to the label (0_1) in case of marked points in φ

Figure 10 shows the cut and paste actions related to the inverted mapping with the pattern $\mathbf{p} = x^2\bar{x}$.

At this point, we can describe the complete mapping defined by the succession rule (15). In particular Figure 11 shows this complete mapping with the pattern $\mathbf{p} = x^2\bar{x}$ and Figure 12 sketches some levels of the generating tree for the paths in $F^{[\mathbf{p}]}$ enumerated according to the number of the rise steps.

This construction generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} instances are coded by a sequence of labels ending by a marked one, say (\bar{k}) , and contain

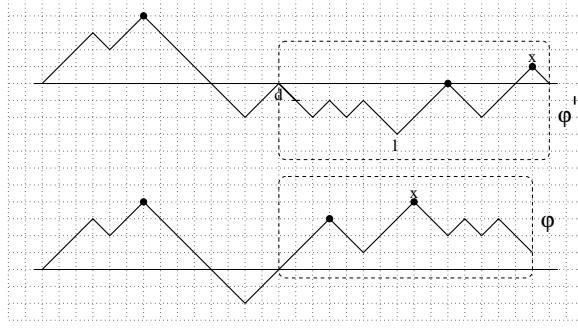


Figure 10: The inverted mapping related to the label (0_1) in case of marked points in φ

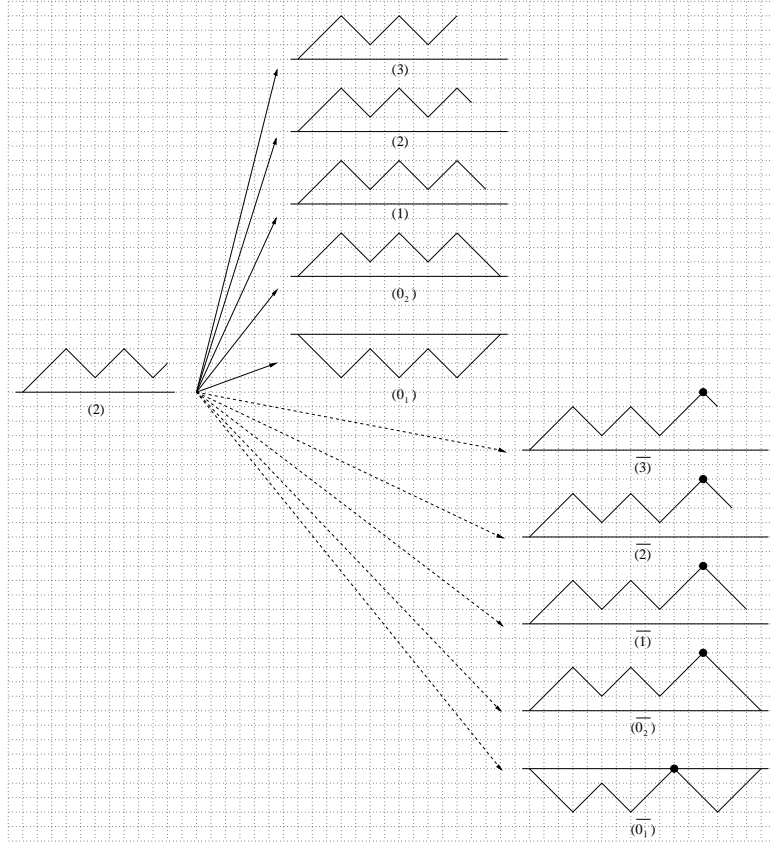
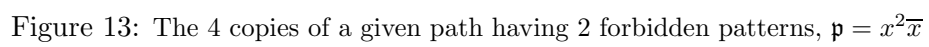
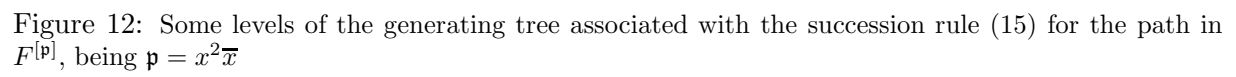


Figure 11: The set of lattice paths obtained from a given (k) , by means of the succession rule (15)

an odd number of marked forbidden patterns, and 2^{C-1} instances are coded by a sequence of labels ending by a non-marked one, say (k) , and contain an even number of marked forbidden patterns. For example, Figure 13 shows the 4 copies of a given path having 2 forbidden patterns $\mathbf{p} = x^2\bar{x}$, where the sequences of labels show the derivation of each path in the generating tree.

This observation is due to the fact that when a path is obtained according to the first production of (15) then no marked forbidden pattern is added. Moreover, when a path is obtained according to the second production of (15) exactly one marked forbidden pattern is added. In any case, the actions performed to obtain the label (0_1) do not change the number of marked forbidden patterns in the path.



Theorem 5.1 *The generating tree of the lattice paths in $F^{[p]}$, where $p = x^{j+1}\bar{x}^j$, according to the number of rise steps, is isomorphic to the tree having its root labelled (0) and recursively defined by the succession rule (15).*

Proof. We have to show that the algorithm described in the previous pages is a construction for the set $F^{[p]}$ according to the number of rise steps. This means that all the paths in F with n rise steps are obtained. Moreover, for each obtained path ω in $F \setminus F^{[p]}$, having C forbidden patterns, with n rise steps and (k) as last label of the associated code, a path ω' in $F \setminus F^{[p]}$ with n rise steps, C forbidden patterns and (\bar{k}) as last label of the associated code is also generated having the same form as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa.

The first assertion is an immediate consequence of the construction according to the first production of (15).

In order to prove the second assertion we have to distinguish two cases (which in their turn are subdivided in 5 and 3 subitems respectively) depending on whether the last forbidden pattern is marked or not. For sake of completeness we report the entire proof, which is indeed rather cumbersome. Anyhow, the interested reader could skip all the subitems, except the first ones. In fact, all the others are obtained from these by means of slight modifications.

We denote by h be the ordinate of the peak of the last forbidden pattern.

First case: the last forbidden pattern in ω is marked. We consider the following subcases: $h > j$, $h = j$, $0 < h < j$, $h = 0$ and $h < 0$.

- 1) $h > j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu x^{j+1}\bar{x}^f \nu$, where $\mu \in F$, $\nu \in F^{[p]}$ and $j \leq f \leq d + j + 1$ where $d \geq 0$ is the ordinate of the endpoint of μ (see Figure 14).

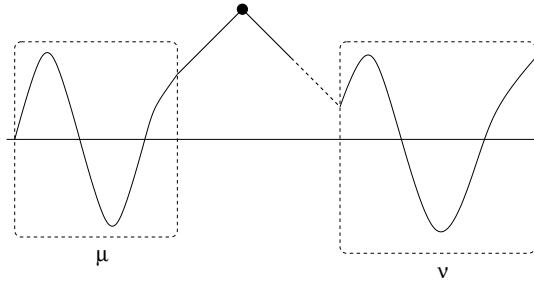


Figure 14: A graphical representation of the path ω in the case $h > j$

The path ω' which kills ω is obtained by performing on μ the following: add the path x^j by applying j times the mapping associated to $(k) \xrightarrow{1} (k+1)$ of the first production of (15), add the path $x\bar{x}^f$ by applying the mapping associated to $(k) \xrightarrow{1} (d+j+1-f)$ of the first production of (15). The path ν in ω' is obtained as in ω .

- 2) $h = j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu\bar{x}\gamma x^{j+1}\bar{x}^j\nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[p]}$ (see Figure 15). We observe that the path γ can contain marked points, with ordinate $b < j$, or not. If the path γ contains no marked point, then it remains strictly under the x -axis, otherwise the marked forbidden patterns intersect the x -axis when $0 \leq b < j$. In the following cases we consider a path γ having the same property.

The path ω' which kills ω is obtained by performing on $\mu\bar{x}\gamma x$ the following: add the path x^{j-1} by applying $j-1$ times the mapping associated to $(k) \xrightarrow{1} (k+1)$ of the first production

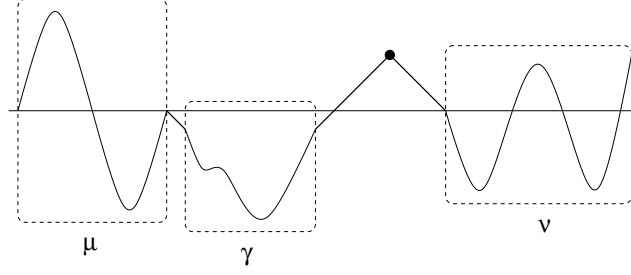


Figure 15: A graphical representation of the path ω in the case $h = j$

of (15), add the path $x\bar{x}^j$ by applying the mapping associated to $(k) \xrightarrow{1} (0_2)$ of the first production of (15). The path ν in ω' is obtained as in ω .

- 3) $0 < h < j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu\bar{x}\gamma x^{j+1}\bar{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[p]}$ (see Figure 16). We observe that the path η remains strictly under the x -axis. In the following cases we consider a path η having the same property.

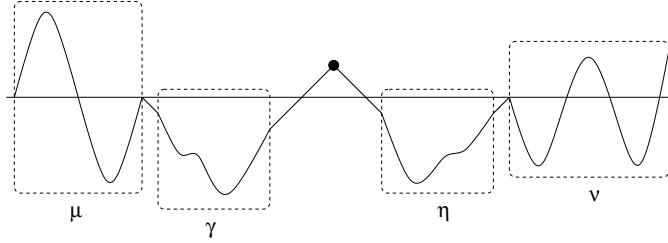


Figure 16: A graphical representation of the path ω in the case $0 < h < j$

The path ω' which kills ω is obtained by performing on $\mu\bar{x}\gamma x^{j+1-h}$ the following: add the path x^{h-1} by applying $h - 1$ times the mapping associated to $(k) \xrightarrow{1} (k + 1)$ of the first production of (15), add the path $x\bar{x}^h$ by applying the mapping associated to $(k) \xrightarrow{1} (0_2)$ of the first production of (15), add the path $\bar{x}^{j-h}\eta x$ by applying consecutive and appropriate mappings of the first production of (15) and these mappings must be completed by performing the actions giving the label (0_1) in case of no marked points. The path ν in ω' is obtained as in ω .

- 4) $h = 0$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu\bar{x}\gamma x^{j+1}\bar{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[p]}$ (see Figure 17).

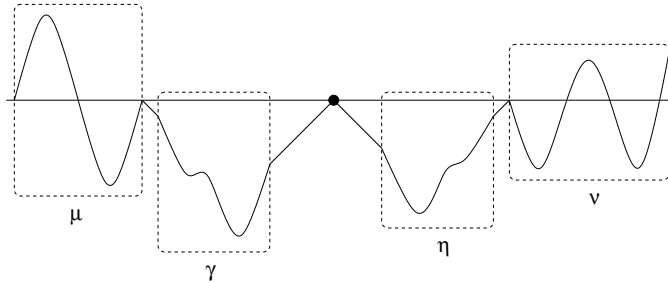


Figure 17: A graphical representation of the path ω in the case $h = 0$

The path ω' which kills ω is obtained by performing on $\mu\bar{x}\gamma x^{j+1}$ the following: add the path $\bar{x}^j\eta x$ by applying consecutive and appropriate mappings of the first production of (15), apply the actions giving the label (0_1) in case of no marked points. The path ν in ω' is obtained as in ω .

- 5) $h < 0$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu\bar{x}\gamma x^{j+1}\bar{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[p]}$ (see Figure 18).

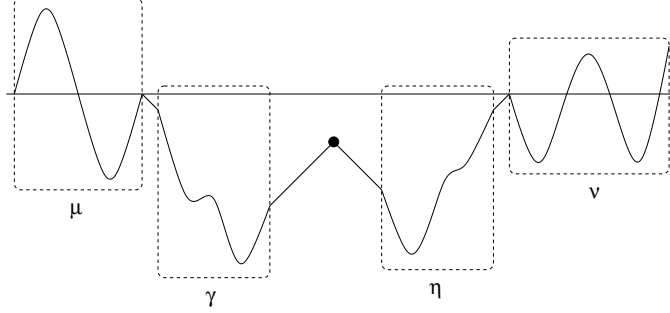


Figure 18: A graphical representation of the path ω in the case $h < 0$

We distinguish two subcases: in the first one the path γ contains no marked points and remains strictly under the x -axis and in the second one the path γ contains at least a marked point.

In the first subcase, the path ω' which kills ω is obtained by performing on μ the following: add the path $\bar{x}\gamma x^{j+1}\bar{x}^j\eta x$ by applying consecutive and appropriate mappings of the first production of (15), apply the actions giving the label (0_1) in case of no marked points. The path ν in ω' is obtained as in ω .

In the second subcase, we consider the rightmost point l of the path $\bar{x}\gamma x^{j+1}\bar{x}^j\eta x$ with lowest ordinate. The path ω' which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\bar{x}^j\eta x$ running from l to the endpoint of the path by applying consecutive and appropriate mappings of the first and second production of (15), add the path in $\gamma x^{j+1}\bar{x}^j\eta x$ running from its initial point to l by applying consecutive and appropriate mappings of the first and second production of (15), apply the cut and paste actions giving the label (0_1) in case of marked points. Obviously, the last forbidden pattern in the path must be generated by applying consecutive and appropriate mappings of the first production of (15). The path ν in ω' is obtained as in ω .

Second case: the last forbidden pattern in ω is not a marked forbidden pattern. We consider the following subcases: $h > j$, $h = j$ and $h < j$.

- 1) $h > j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu x^{j+1}\bar{x}^f\nu$, where $\mu \in F$, $\nu \in F^{[p]}$ and $j \leq f \leq d + j + 1$ where $d \geq 0$ is the ordinate of the endpoint of μ (Figure 19).

The path ω' which kills ω is obtained by performing on μ the following: add the path $x^{j+1}\bar{x}^f$ by applying the mapping associated to $(k) \xrightarrow{j+1} (\overline{d + j + 1 - f})$ of the second production of (15). The path ν in ω' is obtained as in ω .

- 2) $h = j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu\bar{x}\gamma x^{j+1}\bar{x}^j\nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[p]}$ (see Figure 20). We observe that the path γ can contains marked points, with ordinate $b < j$, or not. If the path γ contains no marked point, then it remains strictly under the

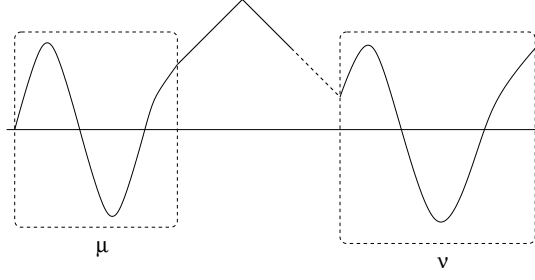


Figure 19: A graphical representation of the path ω in the case $h > j$

x -axis, otherwise the marked forbidden patterns intersect the x -axis when $0 \leq b < j$. In the following case we consider a path γ having the same property.

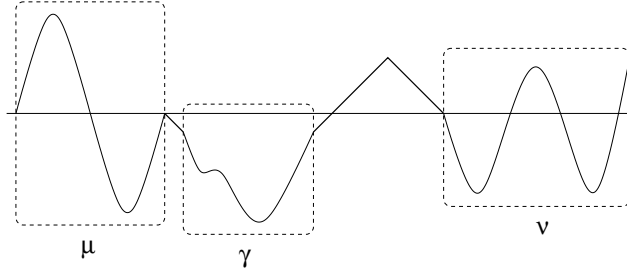


Figure 20: A graphical representation of the path ω in the case $h = j$

Let l be the rightmost point of the path $\bar{x}\gamma x^{j+1}\bar{x}^j$ with lowest ordinate. The path ω' which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\bar{x}^j$ running from l to the endpoint of the path by applying consecutive and appropriate mappings of the first and second production of (15), add the path in $\gamma x^{j+1}\bar{x}^j$ running from its initial point to l by applying consecutive and appropriate mappings of the first and second production of (15), apply the cut and paste actions giving the label (0_1) in case of marked points. Obviously, the last forbidden pattern in the path must be generated by applying the mapping of the second production of (15). The path ν in ω' is obtained as in ω .

- 3) $h < j$: Each path ω in $F \setminus F^{[p]}$ can be written as $\omega = \mu \bar{x} \gamma x^{j+1} \bar{x}^j \eta x \nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[p]}$ (see Figure 21). We observe that the path η remains strictly under the x -axis.

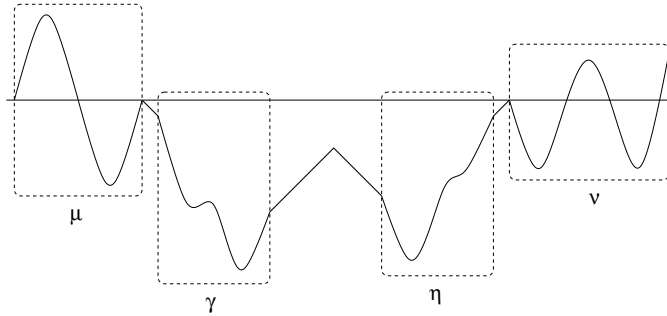


Figure 21: A graphical representation of the path ω in the case $h < j$

Let l be the rightmost point of the path $\bar{x}\gamma x^{j+1}\bar{x}^j \eta x$ with lowest ordinate. The path ω' which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\bar{x}^j \eta x$

running from l to the endpoint of the path by applying consecutive and appropriate mappings of the first and second production of (15), add the path in $\gamma x^{j+1} \bar{x}^j \eta x$ running from its initial point to l by applying consecutive and appropriate mappings of the first and second production of (15), apply the cut and paste actions giving the label (0_1) in case of marked points. Obviously, the last forbidden pattern in the path must be generated by applying the mapping of the second production of (15). The path ν in ω' is obtained as in ω .

We observe that for each path ω in $F \setminus F^{[p]}$, having C forbidden patterns, with n rise steps and last label (k) , there exists one and only one path ω' in $F \setminus F^{[p]}$ with n rise steps, C forbidden patterns and last label (\bar{k}) having the same form as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa.

This assertion is an immediate consequence of the constructions in the proof, since the described actions are univocally determined. Therefore, it is not possible to obtain a path ω' which kills a given path ω applying two distinct procedures. ■

6 Conclusions and further developments

In this paper we study the enumeration, according to the number of ones, of particular binary words excluding a fixed pattern $\mathbf{p} = 1^{j+1}0^j$, $j \geq 1$. Initially, we have solved the problem algebraically by means of Riordan arrays. This approach allows us to obtain a jumping and marked succession rule describing the growth of such words. Note that, it is not possible to associate to a word a path in the generating tree obtained by the succession rule. This problem is solved by means of an algorithm constructing all binary words having a fixed number of ones and eliminating the words which contain the forbidden pattern $\mathbf{p} = 1^{j+1}0^j$, $j \geq 1$.

Further developments could investigate for a unified proof simpler than the one given in this paper. Successive studies should take into consideration binary words avoiding different forbidden patterns both from an enumerative and a constructive point of view. A first step could be the generalization of the forbidden pattern \mathbf{p} , passing from $\mathbf{p} = 1^{j+1}0^j$, $j \geq 1$ to $\mathbf{p} = 1^j0^i$, $0 < i < j$.

Afterwords, it should be interesting to study words avoiding patterns having a different shape, that is not only patterns consisting of a sequence of rise steps followed by a sequence of fall steps. One could also consider forbidden patterns on an arbitrary alphabet and to investigate the properties both of words avoiding that pattern and of the combinatorial objects.

Finally, we could think of studying words avoiding more than one pattern and the related combinatorial objects, considering various parameters.

7 Acknowledgements

The authors wish to thank the anonymous reviews whose comments and suggestions greatly improved the readability of the present paper.

References

- [1] S. Bacchelli, L. Ferrari, R. Pinzani, R. Sprugnoli. Mixed succession rules: The commutative case. *Journal of Combinatorial Theory, Series A*, 117:568-582, 2010.

- [2] D. Baccherini, D. Merlini, and R. Sprugnoli. Binary words excluding a pattern and proper Riordan arrays. *Discrete Mathematics*, 307:1021-1037, 2007.
- [3] D. Baccherini, D. Merlini, and R. Sprugnoli. Level Generating Trees and proper Riordan arrays. *Applicable Analysis and Discrete Mathematics*, 2(1):69-91, 2008.
- [4] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani. ECO: a methodology for the Enumeration of Combinatorial Objects. *Journal of Difference Equations and Applications*, Vol.5 435-490, 1999.
- [5] E. Barcucci, A. Del Lungo, E. Pergola and R. Pinzani, Permutations avoiding an increasing number of length-increasing forbidden subsequences. *Discrete Mathematics and Theoretical Computer Science* 4:31-44, 2000.
- [6] F. R. K. Chung, R. L. Graham, V. E. Hoggatt, M. Kleimann. The number of Baxter permutations. *Journal of Combinatorial Theory, Series A*, 24:382-394, 1978.
- [7] S. Corteel. Séries génératrices exponentielles pour les ECO-systèmes signés. *Proceedings of the 12-th International Conference on Formal Power Series and Algebraic Combinatorics, Moscow, 2000*.
- [8] L. Ferrari, E. Pergola, R. Pinzani, S. Rinaldi. Jumping succession rules and their generating functions. *Discrete Mathematics*, 271 29-50, 2003.
- [9] L. J. Guibas and M. Odlyzko. Long repetitive patterns in random sequences. *Zeitschrift fur Wahrscheinlichkeitstheorie*, 53:241-262, 1980.
- [10] L. J. Guibas and M. Odlyzko. String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory, Series A*, 30:183-208, 1981.
- [11] D. Merlini, D. G. Rogers, R. Sprugnoli, and M. C. Verri. On some alternative characterizations of Riordan arrays. *Canadian Journal of Mathematics*, 49(2):301-320, 1997.
- [12] D. Merlini and R. Sprugnoli. Algebraic aspects of some Riordan arrays related to binary words avoiding a pattern. *In Words2009*, 2009.
- [13] D. Merlini, R. Sprugnoli, M. C. Verri. An Algebra for proper generating tree. In *Algorithms, trees, combinatorics and probabilities*. Trends in Mathematics, Mathematics and Computer Science 127-139, 2000.
- [14] D. Merlini, M. C. Verri. Generating trees and proper Riordan Arrays. *Discrete Mathematics*, 218:167-183, 2003.
- [15] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, Reading, MA, 1996.
- [16] L. W. Shapiro, S. Getu, W. J. Woan and L. Woodson. The Riordan group. *Discrete Applied Mathematics*, 34:229-239, 1991.
- [17] R. P. Stanley. *Enumerative Combinatorics*, volume 2. Cambridge University Press, Cambridge, 1999.
- [18] J. West, Generating trees and the Catalan and Schröder numbers, *Discrete Mathematics*, 146:247-262, 1995.